

## Claims

We claim:

- 1 1. A method for learning a structure of a video to detect events in the video  
2 consistent with the structure, comprising:  
3       selecting sets of features from the video;  
4       updating a hierarchical statistical model for each set of features;  
5       evaluating an information gain of the hierarchical statistical model;  
6       filtering redundant features;  
7       updating the hierarchical statistical model based on the filtered features;  
8       applying a Bayesian information criteria to each model and feature set pair;  
9 and  
10       rank ordering the model and feature set pairs to learn the structure and detect  
11 the events in the video in an unsupervised manner.
- 1 2. The method of claim 1, in which the hierarchical statistical model uses Gaussian  
2 mixtures.
- 1 3. The method of claim 1, in which the hierarchical statistical model uses hidden  
2 Markov models.
- 1 4. The method of claim 3, in which states of events in the video are modeled as  
2 low-level hidden Markov models in the hierarchical hidden Markov model, and the  
3 events are modeled as a high-level Markov chain in the hierarchical hidden  
4 Markov model.

1 5. The method of claim 1, in which the features include dominant color ratios,  
2 motion intensity, a least-square estimates of camera translation, audio volume,  
3 spectral roll-off, low-band energy, high-band energy, zero-crossing rate (ZCR).

1 6. The method of claim 1, in which the features are filtered with a Markov blanket.

1 7. The method of claim 1, in which the evaluating is performed using expectation  
2 maximization and a Markov chain Monte Carlo method.